Veri seti hakkında genel bilgi:

Veri setinde çocukların cinsiyetleri, yaşları, sahip oldukları bilye ve taso sayıları ile çocukların saygı görüp görmedikleri bilgileri yer alıyor. İlk olarak bu verileri genel olarak inceledik ve çocukların cinsiyet, yaş ve oyuncağa sahip olma durumlarına göre dağılımlarını çizdik.

Cinsiyet Dağılımı:

Görselde çocukların cinsiyete göre dağılımını bar grafik olarak gördük. Bu grafik, veri setinde hangi cinsiyetten kaç çocuk olduğunu bize gösterir.

Yaş Dağılımı:

Yaş dağılımını yoğunluk grafiği ile gösterdik. Bu dağılım, çocukların yaşlarına göre ne kadar yoğunlaştığını anlamamızı sağlar.

Oyuncak Dağılımları:

Çocukların sahip oldukları bilye ve taso sayılarına göre histogramlar çizdik. Bu dağılımlar, çocukların hangi oyuncaktan ne kadar sahip olduğunu gösterir.

Saygı Görme Durumu:

Çocukların saygı görme durumunu pasta grafiği ile gösterdik. Bu, çocukların ne kadarının saygı gördüğünü ve ne kadarının saygı görmediğini yüzdelik olarak gösterir.

Model Oluşturma ve Tahmin:

Çocukların yaş ve bilye sayılarına göre saygı görme durumlarını tahmin etmek için bir karar ağacı modeli kullandık. Verileri eğitim ve test verisi olarak böldük ve modeli eğitim verisi ile eğittik. Ardından test verileri üzerinde tahminlerde bulunduk.

Sonuçlar:

Son olarak, modelin performansını ölçmek için bir confusion matrix çizdik. Confusion matrix, modelin doğru ve yanlış tahminlerini gösterir. Bu, modelin ne kadar iyi çalıştığını anlamamıza yardımcı olur.

Bu analiz sonucunda, çocukların yaş ve bilye sayılarına göre saygı görme durumlarını tahmin edebildik. Ancak bu tahminlerin doğruluğu, modelin confusion matrix'inde görülebilir. Daha iyi sonuçlar elde etmek için farklı modeller denenebilir ya da mevcut modelin hiperparametreleri ayarlanabilir. Bu çalışma, çocukların saygı görme durumlarını belirlemek için ilk adımlardan biri olabilir ve daha kapsamlı bir çalışma ile daha sağlıklı sonuçlar elde edilebilir.

Kaynakça : https://github.com/OkanBey11/makine-ogrenmesi-sosyolojik-iliski